

19846  
SEARCH REQUEST FORM

Examiner # (Mandatory): \_\_\_\_\_ Requester's Full Name: \_\_\_\_\_

Art Unit \_\_\_\_\_ Location (Bldg/Room#): \_\_\_\_\_ Phone (circle 305 306 308) \_\_\_\_\_

Serial Number: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

Title of Invention \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Date: \_\_\_\_\_

Keywords (include any known synonyms registry numbers, explanation of initialisms): \_\_\_\_\_

## Search Topic:

Please write detailed statement of the search topic, and the concept of the invention. Describe as specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples of relevant citations, authors, etc., if known. You may include a copy of the abstract and the broadcast or most relevant claim(s).

## STAFF USE ONLY

Searcher: Belcher, C 4994

Searcher Phone #: \_\_\_\_\_

Searcher Location: \_\_\_\_\_

Date Picked Up: \_\_\_\_\_

Date Completed: 09/09-23-99Clerical Prep Time: 5Terminal Time: 20Number of Databases: 1

## Type of Search

\_\_\_\_\_ N.A. Sequence

\_\_\_\_\_ A.A. Sequence

\_\_\_\_\_ Structure (#)

\_\_\_\_\_ Bibliographic

\_\_\_\_\_ Litigation 1

\_\_\_\_\_ Fulltext

\_\_\_\_\_ Procurement

\_\_\_\_\_ Other

## Vendors (include cost where applicable)

\_\_\_\_\_ STN

\_\_\_\_\_ Questel/Orbit

\_\_\_\_\_ Lexis/Nexis

\_\_\_\_\_ WWW/Internet

\_\_\_\_\_ In-house sequence systems (list)

\_\_\_\_\_ Dialog

\_\_\_\_\_ Dr. Link

\_\_\_\_\_ Westlaw

\_\_\_\_\_ Other (specify) Compu Gen

1870

1871

1872

1873

1874

1875

1876

1877

1878

1879

1880

1881

1882

1883

1884

1885

1886

1887

1888

1889

us-08-671-757-1.rge

~~Keith~~ ~~Michael~~ ~~Turner~~, S  
~~09/01/1408~~ 09/01/50 Page 1  
Seg. 1Ds 142

AF082661 Meiotherm  
AF000078  
AF000078 Rhizobium  
AF000078  
U75930 Orygia pseu  
AC004261 Arabidops  
AF000118 Homo sapi  
AF000118 Homo sapi  
AF000165 Homo sapi  
AF000165

AP000315 Homo sapi

5138A5  
AF009899  
Z83818 Human DNA s  
AF009899 Rattus no

AIJ132230	Rattus no
NI66107	Rattus norv

IS796E4  
PNEP  
D11128 Pneumonia v  
AI.022337 Homo sapi

AB017599 Helicobac  
226883 H pylori ge

000/145	007145	Helicobacte
VACA26	Y14742	Helicobacte

VACA3	Y14740	Helicobacte
VACA31	Y14743	Helicobacte

AJ006967 *Helicobac*  
S72494 140 kDa cwt

030402	AF030402	Helicobac
035611	AF035611	Helicobac

049622	AF049622	Helicobac
049627	AF049627	Helicobac

049631	AF049631	Helicobac
049632	AF049632	Helicobac

049638	AF049638	Helicobac
049646	AF049646	Helicobac

049648	AF049648	Helicobac
049650	AF049650	Helicobac

049652	AF049651	Helicobac
049651	AF049652	Helicobac

031813	AF091813	Helicobac
007659	AC007659	Arahidons

AL034421 Homo sapi

LBYV3PH	X53462	Beet yellow
IYORA	A / 34 / 8	Beet yellow

ALV22350 Homo sapi  
Continuation (4 of

Accession	Species
C007185	<i>Drosophila</i>
C007392	<i>Homo sapi</i>

## ALIGNMENTS

DNA	PAT	06-MAR-1998
-----	-----	-------------

250152413.

 $\bar{A}.$ 

action of the pyruvate kinase: production of

8  
 6  
 5  
 4  
 3  
 2  
 1

33596 970116

.....  
fiers

identified"

---

```

BASE COUNT      9 a      5 c      8 g      4 t
ORIGIN

Query Match      100.0%; Score 26; DB 5; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ATGCCTCGAGTCGAAAGCAAGATG 26
        |||||||
Db      1 ATGCCTCGAGTCGAAAGCAAGATG 26

RESULT  2
AF082661/c 2657 bp DNA BCT 11-NOV-1998
LOCUS      Meliothermus ruber proba operon, complete sequence.
ACCESSION  AF082661
VERSION    g3859497
KEYWORDS   AF082661.1 GI:3859497

SOURCE     Meliothermus ruber.
ORGANISM   Bacteria; Thermus/Deinococcus group; Thermus group; Meliothermus.
REFERENCE  1 (bases 1 to 2657)
AUTHORS   Yaklichkin,S.Y., Zimina,M.S., Yurchenko,Y.V., Hromov,I.S. and
           Neumiyakin,L.V.
TITLE      Molecular cloning and sequence analysis of the proA gene from
           thermophilic eubacterium Thermus ruber

JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 2657)
AUTHORS   Yaklichkin,S.Y., Zimina,M.S. and Neumiyakin,L.V.
TITLE      Direct Submission
JOURNAL    Submitted (06-AUG-1998) Center of Biotechnology and Molecular
           Diagnostics, Institute of Molecular Genetics, RASc, Kurchatov
           Square, Moscow 123184, Russia
           3 (bases 1 to 2657)
           Yaklichkin,S.Y., Zimina,M.S. and Neumiyakin,L.V.
           Direct Submission
           Submitted (10-NOV-1998) Center of Biotechnology and Molecular
           Diagnostics, Institute of Molecular Genetics, RASc, Kurchatov
           Square, Moscow 123184, Russia
           Sequence update by submitter
           On Nov 11, 1998 this sequence version replaced gi:3514027.
           Location/Qualifiers
             1..2657
               /organism="Meliothermus ruber"
               /strain="40"
               /db_xref="taxon:277"
             1..2657
               /gene="proA operon"
             61..1233
               /gene="proB"
             61..1233
               /gene="proB"
             /gene="proB"
             /codon_start=1
             /transl_table=11
             /product="gamma glutamyl kinase"
             /protein_id="AAC72812.1"
             /db_xref="PID:g3859498"
             /db_xref="GI:3859498"
             /translation="MRNAPRSVTTTQAMHGRPLTAOTHRRLVYKVSATLSGPOGRQ
             HOLIAAVALRAEGRVIVSSGAQGTGKQTLTEKPSMGKALAVGCPITLM
             LIMEQAFSWYDLKVAQVLLTIEDLAHRRYINARQTLLETLEWGIPIINENDTVME
             EIKFGDNQSLIASIVGADLLILSLIEALYEADPRTHPEAQPPIYVERVAGVLR
             MAGSPNIVGTGKMSKLLAEKAAAGIPIILLPGTRPOSIAEALQGAAGVGLTFAG
             QRRYSGRKTLMLYOLPKPGEEVVDAGAKAAROGASLIPAGLILEVGGQVGEAVVC
             LDEQGNLIGVGLVNTSAAELAKIRKRTREIALGKTNDEAIHRDYPALASELE"
             1239..1244
               /gene="proA operon"
               /note="putative"
               /note="1249..2502"
               /gene="proA"

BASE COUNT      490 a      899 c      866 g      402 t
ORIGIN

Query Match      71.5%; Score 18.6; DB 2; Length 2657;
Best Local Similarity 84.0%; Pred. No. 99;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY      1 ATGCCTCGAGTCGAAAGCAAGAT 25
        |||||||
Db      634 AAGCCTCGAGTCGAGAGCAAGAT 610

RESULT  3
AE000078/c 15586 bp DNA BCT 12-DEC-1997
LOCUS      Rhizobium sp. NGR234 plasmid pNGR234a, section 15 of 46 of the
DEFINITION complete plasmid sequence.
ACCESSION  AE000078 U00090
VERSION    g2182439
KEYWORDS   AE000078.1 GI:2182439

SOURCE     Rhizobium sp. NGR234.
ORGANISM   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
           Rhizobium sp. NGR234
           Rhizobiaceae; Rhizobium.
           1 (bases 1 to 15586)
           Freiberg,C., Fellay,R., Bairoch,A., Broughton,W.J., Rosenthal,A.
           and Perret,X.
           Molecular basis of symbiosis between Rhizobium and legumes
           Nature 387 (6631), 394-401 (1997)
           2 (bases 1 to 15586)
           Freiberg,C.
           Direct Submission
           Submitted (22-NOV-1996) Genome Analysis, Institute for Molecular
           Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
           3 (bases 1 to 15586)
           Freiberg,C.
           Direct Submission
           Submitted (12-DEC-1997) Genome Analysis, Institute for Molecular
           Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
           Update by submitter
           Location/Qualifiers
             1..15586
               /organism="Rhizobium sp. NGR234"
               /plasmid="pNGR234a"
               /strain="NGR234"
               /db_xref="taxon:394"
             <1..>15586
               /organism="Rhizobium sp. NGR234"
               /db_xref="taxon:394"
               /clone="pXB182"
               /complement(82..130)
               /note="nod box-like sequence"
               /note="region: nod box 10"
             misc-feature
               /note="region: nod box 10"

```

```

gene
/evdence=not_experimental
complement(1178..870)
/gene="y41c"
complement(1178..870)
/gene="y41c"
/feature="hypothetical 25.8 kd protein"
/codon_start=1
/transl_table=11
/product="y41c"
/protein_id="AAB91698.1"
/db_xref="PID:92182440"
/db_xref="GI:2182440"
/translation="MYREBAFNAAPASGVLDLAAAYGOOSNGRDEPRFGDIER
LIGTGTGKRLPGLTPTVETAMRNKIPKKEPDLTILLENVLCVRRRVAGIRSK
PASEIMGOVIRALDGPLAPIHCASRAFCDDCNDPVTROSIVLCHREKKAIGYVA
AMLRNASRVELSDVSRRIIAOHQLRISORPRAOYCPASRRNRNTEMGGLRFGII
VRLDRHOSILVN"
complement(812..2800)
/gene="y41d"
complement(812..2800)
/gene="y41d"
/feature="probable monooxygenase"
/codon_start=1
/transl_table=11
/product="y41d"
/protein_id="AAB91699.1"
/db_xref="PID:92182441"
/db_xref="GI:2182441"
/translation="MAHSSDLALFRNRHIAELPSANVPTLLLLYOPTGKEYMLTPPE
LPKSVGNDNGSLAELOTEIRNALDIAICHSAPYKODLSEELIRMLTISE
AEPPEPADMHLKLRRYAGAIIPDPSLIPGPFVLLIGAGMSVAAIRLGLTISY
IOVKKOSTGVMHAAHHYRGGVDTPGHIXSYTPASGNMSTFPPLKEIDYNNRVA
DGEIBSTRTECLVTRYDESLJTHSRKLEPVTETITVNVLSAVGFTTPKP
NLSGIRNEDGVVHTSKMDPEVALDGRKRAVINGASAMQVPAIADVGALTIPOKS
ROMVAPPEKQPVPEPMQELFREVPHYEMLRSLMIDSEVHEALQDPAMPHPD
RSVNAVDRDEATYOTIEGOLAPDLIAVIPSYPFGKMLDNGWRTLLKPHV
TLVGAARVGRSIIHAIDDETHADLIIVASGYDTTLYLPEVIGRNRTYRDWD
DDDCATIGITYAFAFPNFMIXGPTALGHRGSITITIEQIDIVLSYLQMEKRLY
EVECRQDIYOHNRKIOOHOOIMSHSGMSTFPRNDGRVITNSPWLVDYNNLKE
ADLDYRTMPQVDSRLTSGVREGVORPGSRLRRPS"
3544..3951
/gene="y41e"
3544..3951
/gene="y41e"
/feature="hypothetical 15.4 kd protein (possibly
fragmentous); similar to y4rL,y4zA"
/codon_start=1
/transl_table=11
/product="y41e"
/protein_id="AAB91700.1"
/db_xref="PID:92182442"
/db_xref="GI:2182442"
/translation="MOERLALADTTPPAMRSQVGAOPIDLFLASLRTAMKDGATRPD
RPIYAKGRRRPPVLYKATADLRSMFAEPWRTPGSFELSRLAQEYRGDYPDKLLRTL
GRLLKVRSPQADALICTLAKELPQOIAAPH"
3544..3951
/gene="y41f"
/feature="insertion sequence homolog (ISH) 8g: region
containing putative IS element-derived ORF (y41e)"
/region="Region: ISH8g"
complement(4460..5377)
/gene="y41f"
complement(4460..5377)
/gene="y41f"
complement(4460..5377)
/gene="y41f"
/feature="hypothetical 34.1 kd protein"
/codon_start=1
/transl_table=11
/product="y41f"
/protein_id="AAB91701.1"
/db_xref="PID:92182443"
/db_xref="GI:2182443"
/translation="MSGALLMAKKAQRISSALVYGTNNMVDALKLAOKLHGVRDY
VQMHQDRCKRGISVLEHGPYVAFPOKTLPLVIGLGETSCTIVDRSRETLGDI

```

---

```

misc-feature
/feature="insertion sequence homolog (ISH) 11c: region
containing putative IS element-derived ORF (y41g)"
/region="Region: ISH11c"
complement(5839..6111)
/gene="y41g"
complement(5839..6111)
/gene="y41g"
/feature="hypothetical 10.5 kd protein (possibly
fragmentous); similar to y4af and y4bf"
/codon_start=1
/transl_table=11
/product="y41g"
/protein_id="AAB91702.1"
/db_xref="PID:92182444"
/db_xref="GI:2182444"
/translation="MLDLYREHYDGFPTLAEKILIERHLAVSKETLRQIMAGIIV
ALRRERKRREVEQPRGRDCEGLIQIDSGPDTKSTVPLTLSTAR"
6294..7013
/gene="y41h"
6294..7013
/gene="y41h"
/feature="hypothetical 26 kd protein precursor"
/codon_start=1
/transl_table=11
/product="y41h"
/protein_id="AAB91703.1"
/db_xref="PID:92182445"
/db_xref="GI:2182445"
/translation="KRTVAMLLAAVGAVASSTLAVNFCQSKNRNCTGSGFLVD
IFGDSHKLKADFCYVDLVNGVQWQNTKRTIDGASIPPLQDPFGSPMEDGYRAAY
HDWVCDFRVTREKTHREYVDTMASGLERPKALPLFYAVYAGPBMGYLVPEKCA
ACKNCIOMGRKAAFAVOLRGLADQSSAGELKATIKATIDLKERSGDATLDELMAID
EAMPKQTLRDQPRAGDEITK"
complement(7186..9297)
/gene="y41i"
complement(7186..9297)
/gene="y41i"
/feature="hypothetical 76.2 kd integral membrane protein"
/codon_start=1
/transl_table=11
/product="y41i"
/protein_id="AAB91704.1"
/db_xref="PID:92182446"
/db_xref="GI:2182446"
/translation="HRRSTEGTWTVDGSKPTFEIGIAMSQAISGATSAQVDFELIO
ALDAMWKAAEAGAPDLPEYDVRKALSGASAGATTAIGVIAAGREAPATPSPAG
SONIRFTLRLRSWVTSPTLVSPGSDLSLELDAGRPVITLVANNVLTALGAEA
LEATGTLSPRAVASSLHLYMMLNLRVPAIHHNGGOYMMTADHVRHVGGIGT
WKPTSPFPADTDSGTISIAATSLFGAAGASTREPEVLAANMAASGAPIGISPVIA
TATSOYAKAKPPISEDSOELRPIPTVPMAMHVASODYPSFSVSDGLINNDPEVY
RTTLMKDPPEPNERNAEKDRAVITAIAPPEGPPPLDGGEPPLVSLIARVYVATRO
QVRFKPDQLAAVAGETHSRMISPHRVPSPTPGSEEREETISLISGLLGGGVLAE
AFRDHXYQGRNRCQYFLMRHLITIDKNQTLMPGGAERNAVISTKLSDGSMKIDY
PIIPLVGDALPEVPYRVARIDENAFALVYKIEBARLVAVARAVSTETTSARMLGL
NFLLAGRARIYDYLRTLLQELVARDQIEGPIPAALDRPDEVAVIAALLDPAFLD
RTEAGIARTTKIDTSLVREITSLTAGAAGANQVLAWTRIDERSLITVSRRSFL
ATLLQGRSPARLFAKPVYDRK"
7567..15386
/organism="Rhizobium sp. NGR234"
/db_xref="taxon:394"
/clone="BX81661"
complement(9346..11136)
/gene="y41j"
complement(9346..11136)
/gene="y41j"
/feature="hypothetical 65.5 kd protein"
/codon_start=1
/transl_table=11

```



/db\_xref="GI:1911254"  
 /translation="MPCARFVRSLRLTOFEKENVAVAHYDHLMLGLALIDGKVTSAV  
 RREGLSRNVIVSACAAVNVQYAPADDTIMRHOPIYRVCNCHAMADVAPDDHS  
 IARYLACGVALIDHLDVFGFETGEBVELLEVORINAGDL"  
 complement(5669..6151)  
 /note="ORF9: Ptp-2"  
 /codon\_start=1  
 /product="protein tyrosine phosphatase 2"  
 /protein\_id="AAC59008.1"  
 /db\_xref="PID:91911255"  
 /db\_xref="GI:1911255"  
 /translation="MYEASQIDEHVFVGYGNDNEAMLRPIENHAIASISLDSVA  
 PIRHAIPLPGVDHIVCEAPRTCAALPNAMPALYDVRIGKRVIVHCAGASR  
 SAAALVYILMSRQMPYQDALNAVOSKRVAINDFVRFLATCSYRFVDELKPOIV  
 "

CDS  
 complement(6129..6791)  
 /note="ORF10: Ptp-1, AcMNPV ORF1 homolog"  
 /codon\_start=1  
 /product="protein tyrosine phosphatase 1"  
 /protein\_id="AAC59009.1"  
 /db\_xref="PID:91911256"  
 /db\_xref="GI:1911256"  
 /translation="MELSRMAVIYTRTSENGCAGCTGVRRRQVHLPSYRAM  
 PDRMHEYACGAVIEGTRLCFKVPLMAELFEYVTSDEDMTASVLAHSLGAVI  
 DLNTRARYDGAOMYKGLYKRIYVPGRAVDDIYAEFETIDEFERPCFPLAV  
 HHTHGLNRSGLVCRVYMERGVSPDAIARFETARGHKIBRTNYLDDLARKHVRGO  
 PN"

CDS  
 6819..7814  
 /note="ORF11: AcMNPV ORF11 homolog"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AAC59010.1"  
 /db\_xref="PID:91911257"  
 /db\_xref="GI:1911257"  
 /translation="MSLSKLLVAVYGSYNLPDRGSESHIRIYVHEHTNTYVS  
 ASCVRDIATARCNSGLCFVARQLDVSVAARLSAMRCGATGLCADMORALA  
 DIDRNPRLARVGRANIFALDIADISDVYVATNIGIIGRMHPRCSGLARADVF  
 DPDIRAGMWHKFCVLTVMHLYACGAVGASATRLDPAVAKHIGPDNGCAPALAA  
 VYGRCAIGREHFAHKTACMHITFOFMRNDLTPADERPCGVTKDGRCKCTYTD  
 LRTHDALYIHGTTDQKNALEDLDCVNASIDDDCVYVKNRYAQNKKYKA"  
 complement(8180..9142)  
 /note="ORF12: AcMNPV ORF13 homolog"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AAC59011.1"  
 /db\_xref="PID:91911258"  
 /db\_xref="GI:1911258"  
 /translation="MFSWFGWMAADDEVNAEFDEQAVRYAVDQRAHSDLYRDF  
 RCHFEVFRVYVDETAGRCVSVDKGLKISHDLQKNDROHVRQLNELYGAP  
 PAODSLGSLFATKGLHVLQLOFSFANKREYLLAVGAKNDHNDLTKLEAVNFYR  
 KLTNTNSDKFISAKHSFKLEVGARFDEORTLDTKLNALOCAPATETAGVYPR  
 DVTNHPHLAFMKRGVDRGYTOIAPARAGEEHRKRLKEEGEMADVGRAPNDLLA  
 VHCIEEFANGGHRIRLPRKVIYEDCAVNAVDYKKAHLKMT"  
 complement(9094..9825)  
 /note="ORF13: Ief-1, AcMNPV ORF14 homolog"  
 /codon\_start=1  
 /product="late expression factor 1"  
 /protein\_id="AAC59012.1"  
 /db\_xref="PID:91911259"  
 /db\_xref="GI:1911259"  
 /translation="MAPCKYTPERVQMMMDAIVNDSRLAFTDTPRVYVHANFDS  
 AAOLEAVYKNSISDVAVKPLEEGREVMVIDADFKCADAEMLKLVNAGATAMLEF  
 EGRDAVORIMESGNRGHMLKFCGKFKMDAPKSLREHWFNFQPAKLVSGDLIPG  
 SFACVRAVYMTYGDAREDLVLRWPDVDRDVCNANKQIRAPFSYNYKGGYSRCL  
 TQQLQQRKACSAACLAGTPPTSK"  
 9893..11362  
 /note="ORF14: egt, AcMNPV ORF15 homolog"  
 /codon\_start=1  
 /product="ecdysteroid UDP-glucosyltransferase"  
 /protein\_id="AAC59013.1"  
 /db\_xref="PID:91911260"  
 /db\_xref="GI:1911260"  
 /translation="WVFLIALTLATGARAASILAVLPPIPAVSHVYVAYVHALVK

Query Match 70.0% Score 18.2: DB 17: Length 131993:  
 Best Local Similarity 87.0% Pred. No. 1.2e+02;  
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GCCTGAGTGGAAAGCAAGAT 25  
 DB 31011 GCCTGCGCTCAAAAGCAAGAT 31033

RESULT 5  
 LOCUS ATAC004261  
 DEFINITION Arabidopsis thaliana chromosome II BAC T3K9 genomic sequence,  
 complete sequence.  
 ACCESSION AC004261  
 NID 93402695  
 VERSION AC004261.1 GI:3402695  
 KEYWORDS HTG.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core  
 eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 Arabidopsis.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 Rounsley, S.D., Kaul, S., Lin, X., Ketchum, K.A., Crosby, M.L.,  
 Brandon, R.C., Sykes, S.M., Mason, T.M., Kerlavage, A.R., Adams, M.D.,  
 Somerville, C.R. and Venter, J.C.  
 Arabidopsis thaliana chromosome II BAC T3K9 genomic sequence  
 Unpublished  
 2 (bases 1 to 91539)  
 Rounsley, S.D. and Lin, X.  
 Direct Submission  
 Submitted (05-MAR-1998) The Institute for Genomic Research, 9712  
 Medical Center Dr., Rockville, MD 20850, USA, rounsley@tigr.org  
 3 (bases 1 to 91539)  
 Rounsley, S.D.  
 Direct Submission  
 Submitted (07-AUG-1998) The Institute for Genomic Research, 9712  
 Medical Center Dr., Rockville, MD 20850, USA  
 4 (bases 1 to 91539)  
 Lin, X.  
 Direct Submission  
 Submitted (02-FEB-1999) The Institute for Genomic Research, 9712  
 Medical Center Dr., Rockville, MD 20850, USA  
 On Aug 7, 1998 this sequence version replaced gi:334994.  
 Address all correspondence to:  
 Xiaoying Lin  
 The Institute for Genomic Research  
 9712 Medical Center Dr.  
 Rockville, MD 20850, USA  
 e-mail: xlin@tigr.org

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 The orientation of the sequence is from SP6 to T7 end of the BAC  
 clone.  
 Genes were identified by a combination of three methods: Gene  
 prediction programs including GRAIL (available by anonymous ftp  
 from arthur.epm.ornl.gov), GeneFinder (Phil Green, University of  
 Washington), Genscan (Chris Burge,  
 http://www.cbs.dtu.dk/netpene/cbsnetpene.html), and NetplantGene  
 (http://www.cbs.dtu.dk/netpene/cbsnetpene.html), searches of the  
 complete sequence against a peptide database and the Arabidopsis  
 EST database at TIGR (http://www.tigr.org/tigr/at.html).  
 Annotated genes are named to indicate the level of evidence for  
 their annotation. Genes with similarity to other proteins are named





RESULT	6
AP000118	
LOCUS	AP000118
DEFINITION	AP000118 100000 bp DNA PRI 21-MAY-1999
ACCSSION	SICCA3.f444 Homo sapiens genomic DNA of 21q22.1, GART and AML related, SICCA3.f444 region, segment 1/8, complete sequence.
NID	G4730887
VERSION	AP000118.1 GI:4730887
KEYWORDS	HTG.
SOURCE	Homo sapiens DNA.
ORGANISM	Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (sites: ...ances, Catarrhini; Homiidae; Homo.  
Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.  
Homo sapiens 817,199p. Genomic DNA of 21q22.1 GAT and AML region.  
Published Only in Database (1999) In press  
2 (bases 1 to 100000)  
Hirakawa, M., Yamaguchi, H., Imai, K. and Shimada, J.  
Direct Submission  
Submitted (15-APR-1999) to the mainte...

sequence is submitted by Human Genome Sequencing in ALIS project of JST.

Japan Science and Technology Corporation (JST)  
5-3, Yonbancho, Chiyoda-ku, Tokyo 102-0081, Japan

For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.alls.tokyo.jst.go.jp/HGS/>) or send email to [webmaster@www.alls.tokyo.jst.go.jp](mailto:webmaster@www.alls.tokyo.jst.go.jp).

location/Qualifiers

1

**STS**

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
17599..17748
/contig="47S/47A:The location is between each flanking site
of PCR primers."
/db_xref="GDB:191578"
22092..22348
/contig="A005K42:The location is between each flanking site
of PCR primers."
/db_xref="GDB:4570323"
22170..22294
/contig="Cdaa2e11:The location is between each flanking site
of PCR primers."
/db_xref="GDB:441345"
22213..22339
/contig="WI-13874:The location is between each flanking site
of PCR primers."
/db_xref="GDB:4574994"
82240..82341
/contig="SHGC-52061:The location is between each flanking

```

BASE COUNT	ORIGIN
26820 a	19908 c 21667 g 31605 t

QY 1 ATGCTGTGAGGTGGAAGCAAGATG 26  
||||| | ||| | |||||  
Db 85087 ATGCCCCAGGGGTGAAGCAAGATG 85112

RESULT	7	
AP000165		
LOCUS	AP000165	100000 bp DNA
DEFINITION	Homo sapiens genomic DNA, chromosome 21q22.1, D21S22c-AML region, clone B2444F14.T508, segment 1/9, complete sequence.	PRI 14-MAY-1999
ACCESSION	AP000165	
VERSION	g4827130	
KEYWORDS	AP000165.1	GI:4827130
SOURCE	HTG.	
ORGANISM	Homo sapiens	DNA.
	Homo sapiens	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 100000)  
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
Fujiyama,A., Yada,T., Totoki,Y., and Sakaki,Y.  
Homo sapiens 890,291bp genomic DNA of 21q22.1 (REGION: D21S226-AM  
CLONE RANGE: B2344F14-15088)  
Published Only in Database (1999) In press  
2 (bases 1 to 100000)  
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
Fujiyama,A., Yada,T., Totoki,Y., and Sakaki,Y.  
Direct Submission  
Submitted (10-MAY-1999) to the DDB/J/EMBL/GenBank

FEATURES	source
COMMENT	E. coli transposon insertion: the present data does not contain E. coli transposon sequences which integrated in the original/previous sequences. We determined the boundary between the insertion and genomic sequences experimentally, removed the insertion sequences, reconstituted the present data. The sequencing project is supported by Japan Science Technology Corporation (JST) and The Institute of Physical and Chemical Research (RIKEN).
	Location/Qualifiers
	1..100000

BASE COUNT	26797	a	20563	c	21833	g	30807	t
ORIGIN								

Query Match	69.2%	Score 18;	DB 9;	Length 100000;
Best Local Similarity	80.8%;	Pred. No. 1.5e+02;		
Matches 21; Conservative	0;	Mismatches 5;		

QY 1 ATGCTCGAGGTGAAAAGCAAGATG 26  
||||| ||| | |||||  
Db 67753 ATGCCCCAGGCTGTAAGCAAGATG 67778

RESULT 8  
LOCUS AP000315 68109 bp DNA PRI 15-MAY-1999  
DEFINITION Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region, clone:B2344F14, complete sequence.  
ACCESSION AP000315  
NID 94835684  
VERSION AP000315.1 GI:4835684  
KEYWORDS HTG.  
SOURCE Homo sapiens DNA, clone:B2344F14.  
ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eukaryota; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 68109)  
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.  
TITLE Homo sapiens 68,109bp genomic DNA of 21q22.1  
JOURNAL Published only in Database (1999) in press  
REFERENCE 2 (bases 1 to 68109)  
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (13-MAY-1999) to the DDBJ/EMBL/GenBank databases. Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923, Fax:81-42-778-9924)  
COMMENT The sequencing project is supported by Japan Science Technology Corporation (JST) and The Institute of Physical and Chemical Research (RIKEN).  
FEATURES  
source 1. 68109  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/clone="B2344F14"  
/map="21q22.1"  
BASE COUNT 17771 a 13825 c 14920 g 21593 t  
ORIGIN  
Query Match 69.2%; Score 18; DB 9; Length 68109;  
Best Local Similarity 80.8%; Pred. No. 1.6e+02;  
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
OY 1 ATGCTCGAGTCGAAAGCAGATG 26  
Db 67753 ATGCCCCAGGCTGAACGACATG 67778  
|||||  
RESULT 9  
LOCUS AP000316 24518 bp DNA PRI 15-MAY-1999  
DEFINITION Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region, clone:SI85, complete sequence.  
ACCESSION AP000316  
NID 94835685  
VERSION AP000316.1 GI:4835685  
KEYWORDS HTG.  
SOURCE Homo sapiens DNA, clone:SI85.  
ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eukaryota; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 24518)  
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.  
TITLE Homo sapiens 24,518bp genomic DNA of 21q22.1  
JOURNAL Published only in Database (1999) in press

REFERENCE 2 (bases 1 to 24518)  
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (13-MAY-1999) to the DDBJ/EMBL/GenBank databases. Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923, Fax:81-42-778-9924)  
COMMENT The sequencing project is supported by Japan Science Technology Corporation (JST) and The Institute of Physical and Chemical Research (RIKEN).  
FEATURES  
source 1. 24518  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/clone="SI85"  
/map="21q22.1"  
BASE COUNT 7065 a 5274 c 5119 g 7060 t  
ORIGIN  
Query Match 69.2%; Score 18; DB 9; Length 24518;  
Best Local Similarity 80.8%; Pred. No. 1.7e+02;  
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
OY 1 ATGCTCGAGTCGAAAGCAGATG 26  
Db 24407 ATGCCCCAGGCTGAACGACATG 24382  
|||||  
RESULT 10  
LOCUS HS138A5 69013 bp DNA PRI 12-FEB-1997  
DEFINITION Human DNA sequence from PAC 138A5 on chromosome X contains ESTs.  
ACCESSION Z83818  
NID 91841425  
VERSION 283818.1 GI:1841425  
KEYWORDS X.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 69013)  
AUTHORS Wray,P.  
TITLE Direct Submission  
JOURNAL Submitted (05-FEB-1997) Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, UK. E-mail enquires: humuery@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
On Feb 14, 1997 this sequence version replaced gi:1772933.  
de Jong P.J., enquires: http://pacpac.med.buffalo.edu/IMPORANT:  
This sequence is the entire insert of clone 138A5. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unseq' feature key.  
The true left end of clone 138A5 is at 1 in this sequence. The true right end of clone 138A5 is at 69013.  
138A5 is from the human PAC library described in Ioannou A.P. et al Nature Genet 6, 84-89.  
FEATURES  
source 1. 69013  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/map="X"  
/clone="138A5"





ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1162)  
AUTHORS Ni, A., Chai, K.X., Chao, L. and Chao, J.  
TITLE Molecular cloning and expression of rat bradykinin B1 receptor  
JOURNAL Biochim. Biophys. Acta 1442 (2-3), 177-185 (1998)  
MEDLINE 99023739  
REFERENCE 2 (bases 1 to 1162)  
AUTHORS Chai, K.X., Gibbs, T.C., Chao, J. and Chao, L.  
TITLE Direct Submission  
JOURNAL Submitted (05-AUG-1996) Biochemistry and Molecular Biology, Medical  
University of South Carolina, 171 Ashley Avenue, Charleston, SC  
29425-2211, USA

FEATURES  
source location/Qualifiers  
1. 1162  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
1. 1162  
/gene="BKR"  
149. 1162  
/gene="BKR"  
/codon\_start=1  
/product="B1 bradykinin receptor"  
/protein\_id="AAC78505.1"  
/db\_xref="PID:g1762927"  
/db\_xref="GI:1762927"  
/translation="MASEVLELPSPNSRLQAPANTSCSALEMDLLRYLPGFVI  
TICFGLGLNVLVSFLPLPRQMMQORQORLTIAIYLANAASDLVFLGPEVI  
WQGNINRNMFPGLDLCRVASGVIANLNFVSLVVAISODRYRLLYPMTSMGYRR  
RQQAATCLILVAGGLSLIPFLRSRVVDLANSAACILFPEHAMEFARWELNY  
LGLFLVTAIFNFNIIASLRQKASRTRCGSPKSGTTLITIVASPLVCPCPY  
HEFAFDPLVQVAVIQCDSMKETITDGLQLANFPAFVNSCLPLIYFARLRLKTRVL  
GLT"

BASE COUNT 213 a 343 c 312 g 294 t  
ORIGIN

Query Match 69.2%; Score 18; DB 12; Length 1162;  
Best Local Similarity 80.8%; Pred. No. 1.9e+02;  
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 ATGCTCGAGCTCGAAGCAAGATG 26  
Db 770 AAGCCTGTGGGGGAAAGCAAGATG 745

RESULT 14  
LOCUS PNVFP 1657 bp RNA VRL 03-FEB-1999  
DEFINITION Pneumonia virus of mice RNA for fusion protein.  
ACCESSION D11128  
VERSION D11128.1 GI:639678  
KEYWORDS fusion protein; gene 8 (F gene).  
SOURCE Pneumonia virus of mice (strain:15) cell\_line:BSCI RNA.  
ORGANISM Paramyxoviridae; Pneumovirinae; Pneumovirus.  
REFERENCE 1 (bases 1 to 1657)  
AUTHORS Easton, A.J.  
TITLE Direct Submission  
JOURNAL Submitted (13-MAY-1992) to the DDBJ/EMBL/Genbank databases. Andrew  
J. Easton, Warwick University, Department of Biological Sciences,  
Gibbet Hill Road, Coventry, West Midlands, England CV4 7AL, UK  
(Tel:0203-52-3525, Fax:0203-52-3701)  
2 (bases 1 to 1657)  
Chambers, P., Pringle, C.R. and Easton, A.J.  
Sequence analysis of the gene encoding the fusion glycoprotein of  
pneumonia virus of mice suggests possible conserved secondary  
structure elements in paramyxovirus fusion glycoproteins  
J. Gen. Virol. 73 (Pt 7), 1717-1724 (1992)

MEDLINE 92333256  
FEATURES  
source location/Qualifiers  
1. 1657  
/organism="Pneumonia virus of mice"  
/strain="15"  
/db\_xref="taxon:11263"  
/cell\_line="BSCI"  
10. 1623  
/gene="gene 8 (F gene)"  
10. 1623  
/gene="gene 8 (F gene)"  
/codon\_start=1  
/product="fusion protein"  
/protein\_id="BAA01902.1"  
/db\_xref="PID:d1002379"  
/db\_xref="PID:g1468965"  
/db\_xref="GI:1468965"  
/translation="MIGRIFLVLLVIFNFKPIHNTLTETKYSSTGVSFAGYKSL  
RTGMHMTVASIKLSQINIESCKSSNSLALHBLAIVSSAVDELRTLSMALSKRRKRF  
LGLILGLAAVTAGALAKTQVLESELIADPARNNEAVSLTMSVLAKVVDL  
KNEISKELPKINRVSCDHDITAVIRPOOLNKLIVSEFSSNAGLTHTVSEFMT  
DRELSTVIGMAVSAGOKETLSSKATMRNGALISSVADLVVYIQLPFGVMDT  
DCWVIRSIDCHNADRYACIARADNGWYCHNAGSLSPSPDCEIHNGYACDYLK  
SLVPAVIRSECNMNTYNTYDCKISTKTYSTAVLTMCGLVSCYKGNCSCTYINDK  
GIITLDPGCHYISNKGVDROYQNTYVYLSKEGSIYRGPIVLYKPLSPDDK  
FDVAIRDVHISINOTRTEFRASDGLDLSREBNKNNKSYILITLLFVYMLITIMAV  
IGFLYKVLKMRIRNKLKSKSTPGITVLS"

BASE COUNT 528 a 307 c 346 g 476 t  
ORIGIN

Query Match 69.2%; Score 18; DB 17; Length 1657;  
Best Local Similarity 80.8%; Pred. No. 1.9e+02;  
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 ATGCTCGAGCTCGAAGCAAGATG 26  
Db 284 ATGCTTGAAGTCCAAAGCAAGAG 309

RESULT 15  
LOCUS HS796E4  
DEFINITION Homo sapiens chromosome 22 clone 796E4, WORKING DRAFT SEQUENCE, in  
unoriented pieces.  
ACCESSION AL022337  
VERSION g4914536  
KEYWORDS AL022337.10 GI:4914536  
HTG: HTGS\_PHASE1.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 135437)  
AUTHORS Skuue, C.  
TITLE Direct Submission  
JOURNAL Submitted (27-MAY-1999) Wellcome Trust Genome Campus, Hinxton,  
Cambridge, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk  
On May 28, 1999 this sequence version replaced g14902693.  
IMPORTANT: This sequence is unfinished and does not necessarily  
represent the correct sequence. Work on the sequence is in progress  
and the release of this data is based on the understanding that the  
sequence may change as work continues. The sequence may be  
contaminated with foreign sequence from E. coli, yeast, vector,  
phage etc. Order of segments is not known. 800 n's separate  
segments. Unfinished: bk796E4 Contig\_ID: 02320 acc-AL022337  
Length: 11508 bp Unfinished: bk796E4 Contig\_ID: 02324  
acc-AL022337 Length: 3567 bp Unfinished: bk796E4 Contig\_ID:  
02570 acc-AL022337 Length: 86652 bp.  
NOTE: This is a 'working draft' sequence.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will

Thu Sep 23 10:37:29 1999

us-08-671-757-1.rge

Page 12

\* be preserved.  
FEATURES location/Qualifiers

SOURCE

1. .135437  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/clone="796E4"

BASE COUNT 37381 a 28828 c 28389 g 39239 t 1600 others  
ORIGIN

Query Match

69.2%; Score 18; DB 34; Length 135437;

Best Local Similarity 80.8%; Pred. No. 1.5e+02;

Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 ATGCCTCGAGGTCGAAAGCAAGATG 26  
|||||

Db 21136 ATGCCTGTGTGTCTCAAGCAAGATG 21161

Search completed: September 22, 1999, 17:01:57  
Job time: 496 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 1999, 16:53:41 ; Search time 51.18 Seconds

(Without alignments)  
127.100 Million cell updates/sec

Title: US-08-671-757-1

Perfect score: 26  
Sequence: 1 ATGCTCGAGGTGCAAAAGCAAGATG 26

Scoring table: IDENTITY\_NNC

Searched: 311585 seqs, 125096042 residues

Database: N\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	26	100.0	26	1	T62334	Helicobacter pylori
2	18.2	70.0	110000	1	V30458-1	Continuation (2 of
3	18.2	70.0	110000	1	V30459-1	Continuation (2 of
4	17.6	67.7	3960	1	O48732	CT gene. Recombina
5	17.2	66.2	98	1	O81918	Interferon gamma b
6	16.4	63.1	876	1	T61605	Humicola lanuginos
7	16.4	63.1	876	1	T61606	Humicola lanuginos
8	16.4	63.1	876	1	T61593	Humicola lanuginos
9	16.4	63.1	876	1	T61594	Humicola lanuginos
10	16.4	63.1	31	1	V72952	Rat Munc13-1 PCR p
11	16.4	63.1	22191	1	X20600	Polynucleotide seq
12	16.2	62.3	2163	1	T95753	Arabidopsis SCAREC
13	16	61.5	3649	1	O04476	Sequence derived b
14	16	61.5	1657	1	N81265	Gene encoding Chla
15	16	61.5	99	1	N00005	Sequence encoding
16	16	61.5	2518	1	T31256	Hydantoinase codin
17	16	61.5	3190	1	T31257	Continuation (6 of
18	16	61.5	80073	1	T58840-5	Glutamyl-tRNA(Gln)
19	16	61.5	1467	1	X20935	Polynucleotide seq
20	16	61.5	240	1	X17999	DNA encoding a RNA
21	16	61.5	1467	1	X16047	Human kidney amino
22	15.8	60.8	37	1	V19310	Rat p26 gene RACE
23	15.8	60.8	50000	1	X23517	Portion of plasmid
24	15.6	60.0	419	1	O03280	Oct-2 clone encodi
25	15.6	60.0	1714	1	N90718	Pass-3 clone encod
26	15.6	60.0	1847	1	N90719	Plasmid pR3. Immun
27	15.6	60.0	5393	1	O34611	Sequence of plasm
28	15.6	60.0	4641	1	O40294	Sequence of plasm
29	15.6	60.0	6811	1	O41004	Sequence of plasm
30	15.6	60.0	5472	1	O41005	Sequence of plasm
31	15.6	60.0	1133	1	O40246	Sequence of a DNA
32	15.6	60.0	1133	1	O40247	Sequence of a DNA
33	15.6	60.0	1209	1	O40264	Sequence inserted
34	15.6	60.0	4145	1	O40279	Sequence of clone
35	15.6	60.0	4277	1	O40280	Sequence of clone
36	15.6	60.0	6926	1	O40281	Sequence of clone
37	15.6	60.0	32	1	O93242	Plasmid todic2BA
38	15.6	60.0	9020	1	T07652	PS011 FlpV N. Reco
39	15.6	60.0	8710	1	O89650	PS011 FlpV N. Reco
40	15.6	60.0	613	1	T01289	Rhizobium species
41	15.6	60.0	900	1	T47699	Banana polyphenol
42	15.6	60.0	28984	1	V15826	Genomic DNA for in
43	15.6	60.0	4185	1	V43563	Human blue-light p

44 15.6 60.0 9767 1 X13067  
45 15.6 60.0 6560 1 X26118

## ALIGNMENTS

RESULT 1  
ID T62334  
AC T62334;  
DE 23-SEP-1997 (first entry)  
DE Helicobacter pylori flbA gene PCR primer. OLf1A-1.  
KW Flagellum; motility; monoclonal antibody; vaccine; infection; ulcer;  
KW ulcerative colitis; chronic gastritis; duodenal; immunise;  
KW prevention; hook protein; attenuated urease; primer; probe; PCR;  
KW Helicobacter pylori.  
OS EP-752473-A2.  
PD 08-JAN-1997.  
PF 28-JUN-1996; 401445.  
PR 04-JUL-1995; RR-008068.  
PA (INRM) INST NAT SANTE & RECH MEDICALE.  
PI (INSP) INST PASTEUR.  
PI Labigne A, Suerbaum S;  
DR WPI: 97-067450/07.  
PT Helicobacter pylori flbA gene and related protein - regulates  
PT biosynthesis of flagella, useful in compns. and vaccines for  
PS detecting and preventing H. pylori infections  
PS Claim 1; Page 38; 58pp; French.  
CC T62334 and T62335 are PCR primers/probes used for the isolation of the  
CC Helicobacter flbA gene. The flbA gene regulates the biosynthesis of  
CC H. pylori flagellum proteins. The flbA gene may be mutated to produce  
CC a strain of H. pylori having an attenuated phenotype. H. pylori  
CC strains can also be engineered to produce an attenuated urease and  
CC defective (or absent) hook protein. Alternatively the flbA protein can  
CC be used to raise monoclonal antibodies or polyclonal serum for  
CC detecting H. pylori infections. Vaccines especially also contain H. pylori antigens  
CC H. pylori infections. Vaccines especially also contain H. pylori antigens  
CC encoded by the urea, ureb, urec or ured genes.  
SQ Sequence 26 BP; 9 A; 5 C; 8 G; 4 T;

Query Match 100.0%; Score 26; DB 1; Length 26;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCTCGAGGTGCAAAAGCAAGATG 26  
DB 1 ATGCTCGAGGTGCAAAAGCAAGATG 26

RESULT 2  
V30458-1/c  
Continuation (2 of 6) of V30458 from base 100001 (Rhizobium species plasmid pNGR234a.  
WP Sequence split into 6 fragments LOCUS V30458 Accession V30458  
WP Fragment Name Begin End  
WP V30458-0 1 110000  
WP V30458-1 100001 210000  
WP V30458-2 200001 310000  
WP V30458-3 300001 410000  
WP V30458-4 400001 510000  
WP V30458-5 500001 534720

Query Match 70.0%; Score 18.2; DB 1; Length 110000;  
Best Local Similarity 87.0%; Pred. No. 18;  
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 TGCCTCGAGGTGCAAAAGCAAGA 24  
DB 72074 TACATCGAGGTGCAAAAGCAAGA 72052





PT Preparing polypeptide variants with improved functional properties -  
PT by in vivo recombination between opened plasmid and homologous DNA,  
PT to produce e.g. enzymes with improved washing and dishwashing  
PT properties  
PS Example 1; Fig 1; 68pp; English.  
CC A DNA sequence (T61605) comprises the Humicola lanuginosa wild-type  
CC lipase gene in Saccharomyces cerevisiae expression plasmid pJS026.  
CC It encodes a polypeptide (W13560) of 291 amino acid residues.  
CC Plasmid pJS026 has been used to demonstrate an improved method of  
CC preparing positive polypeptide variants. This involves shuffling  
CC homologous DNA sequences in an iterative in vivo recombination  
CC system using a eukaryotic cell (such as yeast) as a recombination  
CC host cell. In an example, the lipase gene from pJS026 was  
CC transformed into S. cerevisiae YNG318 together with a plasmid,  
CC pJS037 (see also T61606), which carries a synthetic H. lanuginosa  
CC lipase gene and which had been opened at approx. the middle of the  
CC lipase coding sequence by digestion with NruI, PstI, and NruI and  
CC PstI. 50% of transformants contained recombinant DNA sequences.  
SQ Sequence 876 BP; 208 A; 229 C; 225 G; 214 T;

Query Match 63.1%; Score 16.4; DB 1; Length 876;  
Best Local Similarity 76.9%; Pred. No. 74;  
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 ATGCGCTGAGGTGCAAAAGCAAGATG 26  
Db 186 ATGCCCGAGGTAGAGAGCGGATG 211  
||||| ||||| ||||| |||||

RESULT 7  
T61606  
ID' T61606 standard; DNA; 876 BP.  
AC T61606;  
DE 21-OCT-1997 (first entry)  
DE Humicola lanuginosa lipase synthetic gene in pJS037.  
KW Lipase; polypeptide variant; in vivo recombination; shuffling;  
KW Saccharomyces cerevisiae; detergent; pJS037; ss.  
OS Synthetic.  
PN WO9707205-A1.  
PD 27-FEB-1997.  
PF 12-AUG-1996; DK0344.  
PR 20-SEP-1995; DK-001047.  
PR 11-AUG-1995; DK-000907.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Okkels JS;  
DR WPI: 97-165290/15.  
DR P-PSDB; W13560.  
PT Preparing polypeptide variants with improved functional properties -  
PT by in vivo recombination between opened plasmid and homologous DNA,  
PT to produce e.g. enzymes with improved washing and dishwashing  
PT properties  
PS Example 1; Fig 2; 68pp; English.  
CC A DNA sequence (T61605) comprises a Humicola lanuginosa synthetic  
CC lipase gene in Saccharomyces cerevisiae expression plasmid pJS037.  
CC It contains 12 additional restriction sites not found in the wild-  
CC type gene (see also T61605). Plasmid pJS037 has been used to  
CC demonstrate an improved method for preparing positive polypeptide  
CC variants. This involves shuffling homologous DNA sequences in an  
CC iterative in vivo recombination system using a eukaryotic cell  
CC (such as yeast) as a recombination host cell. In an example, in  
CC vivo recombination of H. lanuginosa lipase variants was performed  
CC using NruI-cut pJS037 and S. cerevisiae YNG318 as host cell. 2  
CC Transformants (see also W13561-62) were obt'd. that showed improved  
CC wash performance.  
SQ Sequence 876 BP; 205 A; 231 C; 224 G; 216 T;

Query Match 63.1%; Score 16.4; DB 1; Length 876;  
Best Local Similarity 76.9%; Pred. No. 74;  
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 ATGCGCTGAGGTGCAAAAGCAAGATG 26

Db ||||| ||||| ||||| |||||  
186 ATGCCCGAGGTAGAGAGCGGATG 211

RESULT 8  
T61593  
ID T61593 standard; DNA; 876 BP.  
AC T61593;  
DE 21-OCT-1997 (first entry)  
DE Humicola lanuginosa lipase gene in vector pJS026.  
KW Lipase; polypeptide variant; in vivo recombination; shuffling;  
KW Saccharomyces cerevisiae; detergent; pJS026; ss.  
OS Humicola lanuginosa DSM 4109.  
PN WO9707205-A1.  
PD 27-FEB-1997.  
PF 12-AUG-1996; DK0343.  
PR 20-SEP-1995; DK-001047.  
PR 11-AUG-1995; DK-000907.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Okkels JS;  
DR WPI: 97-165289/15.  
DR P-PSDB; W13556.  
PT Preparing polypeptide variants with improved functional properties -  
PT by in vivo recombination between opened plasmid and homologous DNA,  
PT to produce e.g. enzymes with improved washing and dishwashing  
PT properties  
PS Example 1; Fig 1; 68pp; English.  
CC A DNA sequence (T61593) comprises the Humicola lanuginosa wild-type  
CC lipase gene in Saccharomyces cerevisiae expression plasmid pJS026.  
CC It encodes a polypeptide (W13556) of 291 amino acid residues.  
CC Plasmid pJS026 has been used to demonstrate an improved method of  
CC preparing positive polypeptide variants. This involves shuffling  
CC homologous DNA sequences in an iterative in vivo recombination  
CC system using a eukaryotic cell (such as yeast) as a recombination  
CC host cell. In an example, the lipase gene from pJS026 was  
CC transformed into S. cerevisiae YNG318 together with a plasmid,  
CC pJS037 (see also T61594), which carries a synthetic H. lanuginosa  
CC lipase gene and which had been opened at approx. the middle of the  
CC lipase coding sequence by digestion with NruI, PstI, and NruI and  
CC PstI. 50% of transformants contained recombinant DNA sequences.  
SQ Sequence 876 BP; 208 A; 229 C; 225 G; 214 T;

Query Match 63.1%; Score 16.4; DB 1; Length 876;  
Best Local Similarity 76.9%; Pred. No. 74;  
Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 ATGCGCTGAGGTGCAAAAGCAAGATG 26  
Db 186 ATGCCCGAGGTAGAGAGCGGATG 211  
||||| ||||| ||||| |||||

RESULT 9  
T61594  
ID T61594 standard; DNA; 876 BP.  
AC T61594;  
DE 21-OCT-1997 (first entry)  
DE Humicola lanuginosa lipase synthetic gene in pJS037.  
KW Lipase; polypeptide variant; in vivo recombination; shuffling;  
KW Saccharomyces cerevisiae; detergent; pJS037; ss.  
OS Synthetic.  
PN WO9707205-A1.  
PD 27-FEB-1997.  
PF 12-AUG-1996; DK0343.  
PR 20-SEP-1995; DK-001047.  
PR 11-AUG-1995; DK-000907.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Okkels JS;  
DR WPI: 97-165289/15.  
DR P-PSDB; W13556.  
PT Preparing polypeptide variants with improved functional properties -  
PT by in vivo recombination between opened plasmid and homologous DNA,  
PT to produce e.g. enzymes with improved washing and dishwashing

PT Properties  
 CC Example 1; Fig 2. 68pp; English.  
 CC A DNA sequence (T61594) comprises a Humicola lanuginosa synthetic  
 CC lipase gene in Saccharomyces cerevisiae expression plasmid pJS037.  
 CC It contains 12 additional restriction sites not found in the wild-  
 CC type gene (see also T61593). Plasmid pJS037 has been used to  
 CC demonstrate an improved method for preparing positive polypeptide  
 CC variants. This involves shuffling homologous DNA sequences in an  
 CC iterative in vivo recombination system using a eukaryotic cell  
 CC (such as yeast) as a recombination host cell. In an example,  
 CC in vivo recombination of H. lanuginosa lipase variants was  
 CC performed using NuiI-cut pJS037 and S. cerevisiae YN318 as  
 CC host cell. 2 transformants (see also W13557-58) were obtd. that  
 CC showed improved wash performance.  
 SQ Sequence 876 BP; 205 A; 231 C; 224 G; 216 T;

Query Match 63.1%; Score 16.4; DB 1; Length 876;  
 Best Local Similarity 76.9%; Pred. No. 74;  
 Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 ATCCCTCGAGTCGAAAGCAAGATG 26  
 ||||| ||||| ||||| ||||| |||||  
 Db 186 ATCCCTCGAGTCGAGAGGCGGATG 211

## RESULT 10

W72952  
 ID V72952 standard; DNA: 31 BP.

AC V72952;

DT 04-MAR-1999 (first entry)

DE Rat Munc13-1 PCR primer SEQ ID NO:11.

KW Munc13; Doc2-alpha; interacting domain; screening; agonist; antagonist;  
 calcium ion dependent secretion inhibitor; neurotransmitter; hormone;

KW fusion protein; nervous disease; PCR primer; ss.

OS Synthetic.

OS Rattus sp.

PN J10313866-A.

PD 02-DEC-1998.

PF 15-MAY-1997; 126118.

PR 15-MAY-1997; JP-126118.

PT (SH10) SHIONOGI & CO LTD.

DR WPI; 99-074148/07.

PT Screening for agonists or antagonists of binding between Doc2-alpha  
 and Munc13 - used to treat diseases of the nervous system

CS Example 3; Page 27; 33pp; Japanese.

CC The present invention describes a method of screening for agonists or  
 antagonists of the binding between Doc2-alpha and Munc13. The method

CC comprises reacting Doc2-alpha or its homologue with Munc13 or its  
 homologue optionally in the presence of a test substance and selecting

CC the substances which increase or decrease binding. Also described are:  
 CC (1) an agonist or antagonist of the binding between Doc2-alpha and  
 CC Munc13 selected by the above method; (2) a vector expressing Doc2-alpha  
 CC or its homologue used for inhibiting Ca ion-dependent secretion of a  
 CC neurotransmitter or hormone; (3) a vector expressing Munc13 or its  
 CC homologue used for inhibiting Ca ion-dependent secretion of a  
 CC neurotransmitter or hormone; (4) a fusion protein between Doc2-alpha or  
 CC its homologue and a carrier protein; (5) a polypeptide containing  
 CC or its homologue and a carrier protein; (6) a polypeptide containing  
 CC amino acids 13-37 of the sequence of Doc2-alpha, which binds with Munc13  
 CC and comprises at most 90 amino acids; and (7) a polypeptide containing  
 CC amino acids 851-1461 of the sequence of Munc13, which binds with Doc2-  
 CC alpha and comprises at most 904 amino acids. The agonist or antagonist  
 CC can be used to treat diseases of the nervous system. The present  
 CC sequence represents a PCR primer for rat Munc13-1.

SQ Sequence 31 BP; 10 A; 5 C; 11 G; 5 T;

Query Match 63.1%; Score 16.4; DB 1; Length 31;  
 Best Local Similarity 76.9%; Pred. No. 52;  
 Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 ATGCCTCGAGTCGAAAGCAAGATG 26

Db 6 ATTCCTCGAGGCGGACAGAGATG 31  
 || ||||| || ||||| |||||

## RESULT 11

X20600/c  
 ID X20600 standard; DNA: 22191 BP.

AC X20600;

DT 05-MAY-1999 (first entry)

DE Polynucleotide sequence from the genome of Treponema pallidum.

KW Treponema pallidum infection; syphilis; Borrelia infection; animal;  
 enzyme production; ds.

OS Treponema pallidum.

PN WO9859034-A2.

PD 30-DEC-1998.

PF 23-JUN-1997; US-050667.

PT (HUMA-) HUMAN GENOME SCI INC.

DR Fraser CM;

DT WPI; 99-081273/07.

PT New isolated Treponema pallidum nucleic acids - used to develop  
 products for the detection, diagnosis, characterisation, prevention

PT and therapy of T. pallidum infections, particularly syphilis

PS Claim 1; Page 673-685; 1150pp; English.

CC X20300-21243 represent polynucleotide sequences from the genome of

CC Treponema pallidum. The sequences can be used for detection,  
 CC diagnosis, characterisation, prevention and therapy for T. pallidum

CC infections, particularly syphilis. They can also be used for detecting

CC diseases related to Borrelia infections in animals, and for the

CC production of biosynthetic products such as enzymes.

SQ Sequence 22191 BP; 4744 A; 5303 C; 6215 G; 5916 T;

Query Match 63.1%; Score 16.4; DB 1; Length 22191;  
 Best Local Similarity 76.9%; Pred. No. 1e+02;  
 Matches 20; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 ATGCCTCGAGTCGAAAGCAAGATG 26  
 ||||| ||||| ||||| ||||| |||||

Db 15713 AGCCTCGGCTCAAGACCAAGATG 15688

## RESULT 12

T95753/c  
 ID T95753 standard; DNA: 2163 BP.

AC T95753;

DT 11-MAY-1998 (first entry)

DE Arabidopsis SCARECROW gene.

KW SCARECROW; SCR gene; transgenic plant; root; gravitropism;  
 crop improvement; ds.

OS Arabidopsis thaliana.

FS Key Location/Qualifiers

FT CDS 49..2125

FT /tag= a

FT /note= "contains an intron"

FT exon 49..1580

FT /tag= b

FT /number= 1

FT intron 1581..1695

FT /tag= c

FT /number= 1

FT exon 1696..2125

FT /tag= d

FT /number= 2

PN WO9741152-A1.

PD 06-NOV-1997.

PF 25-APR-1997; U07022.

PR 24-APR-1997; US-842445.

PR 26-APR-1996; US-638617.

PA (UNYX) UNIV NEW YORK STATE.

PI Benley PN, Dilaurenzio L, Helariutta Y, Malamy JE,  
 PI Pysch L, Wysocka-Diller J;  
 WPI; 97-549683/50.

DR P-PSDB; W38178.  
 PT DNA encoding Arabidopsis SCARECROW protein - useful to modify plant  
 cell division and therefore alter root development, or alter plant  
 stem or hypocotyl gravitropism  
 PS Claim 6: Page 90-92; 221pp; English.  
 CC This sequence comprises the Arabidopsis SCARECROW (SCR gene)  
 CC genomic region. It was obtained following isolation of an scr-1  
 CC mutant from a T-DNA transformed Arabidopsis seedling that had  
 CC greatly reduced root length. Plant DNA flanking the insertion site  
 CC was isolated from scr-1 by plasmid rescue, and used to isolate the  
 CC wild-type genomic DNA. The SCR gene encodes a 653-amino acid  
 CC protein (see W38178). It is expressed specifically in embryo root  
 CC progenitor tissue and in certain root and stem tissues. It  
 CC regulates a specific asymmetric division, and controls gravitropic  
 CC response in aerial structures and root formation. Nucleic acid  
 CC W38178-201, SCR proteins lacking 1 to 4 of MOTIFS I to VI, or  
 CC SCR MOTIF I, II, III, IV, V or VI are claimed. Transgenic plants  
 CC can be engineered to overexpress the SCR protein, so that cell  
 CC division is increased in roots, resulting in thicker root  
 CC development, while a plant with an altered stem or hypocotyl  
 CC gravitropism is less susceptible to lodging. Plants that contain  
 CC an antisense molecule that suppresses the expression of endogenous  
 CC SCR gene product have thinner root development. A gene of interest  
 CC can be placed under control of a SCR promoter and expressed in a  
 CC plant to confer herbicide, salt, pathogen or insect resistance, or  
 CC when expressed in stems to increase starch, lignin or cellulose  
 CC biosynthesis (all claimed).  
 SQ Sequence 2163 BP; 571 A; 546 C; 439 G; 607 T;

Query Match  
 Best Local Similarity 62.3%; Score 16.2; DB 1; Length 2163;  
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCTCGAGTGTGAAAGCAAGA 24  
 DB 1328 CCTCGAGTGTGAAAGCAAGA 1308

RESULT 13  
 ID 004476  
 AC 004476; standard; cDNA; 3649 BP.  
 DT 05-OCT-1990 (first entry)  
 DE Sequence derived by reverse transcription of RNA1 of AMV.  
 KW Alfalfa mosaic virus; AMV; cryptic infection; probe; ss.  
 OS Nicotiana tabacum L. "Xanthi-nc"  
 FH Key Location/Qualifiers  
 FT cds 100..3486  
 FT /\*tag= a  
 FT /\*note="protein encoded by RNA1"  
 PN US4921802-A.  
 PD 01-MAY-1990.  
 PF 13-JAN-1988; 144692.  
 PR 05-MAY-1983; US-492582.  
 PR 13-JAN-1988; US-144692.  
 PA (PRO) Pioneer Hi-Bred International Ltd.  
 PI Hall TC, Loesch-Fries SL, Jarvis NP, Barker RF;  
 DR WPI; 90-171297/22.  
 DR P-PSDB; R05107.  
 PT Recombinant cDNA plasmids derived from alfalfa mosaic virus RNA -  
 PT for detection of cryptic infection by alfalfa mosaic virus.  
 CC Disclosure: p; English.  
 CC The sequence is that of the bottom (B) - or "RNA1" genome fragment  
 CC of AMV which is necessary to initiate infection. The sequence can  
 CC be used as a probe to detect cryptic infection of plants by AMV.  
 CC See also 004477 and 78.  
 SQ Sequence 3649 BP; 1039 A; 744 C; 808 G; 1058 T;

Query Match  
 Best Local Similarity 61.5%; Score 16; DB 1; Length 3649;  
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 GCCTCGAGGCGAAAGCAAGATG 26  
 DB 182 GACCGCTAGTCGAAAGCAAGCTG 205

RESULT 14  
 ID N81265/c  
 AC N81265; standard; DNA; 1657 BP.  
 DT 24-OCT-1990 (first entry)  
 DE Gene encoding Chlamydial elementary body protein, chlamectin.  
 KW Chlamectin; elementary body; binding protein; vaccine; ss.  
 OS Chlamydia trachomatis serovar L2 (L2/Bu/434).  
 FH Key Location/Qualifiers  
 FT rbs 146..149  
 FT /\*tag= a  
 FT /\*tag= b  
 FT /\*tag= c  
 FT prod-chlamectin  
 FT note="18 kD binding protein"  
 FT /\*tag= c  
 FT signal\_peptide  
 FT 160..172  
 PN EP-293079-A.  
 PD 30-NOV-1988.  
 PF 21-APR-1988; 303584.  
 PR 24-APR-1987; GB-009746.  
 PA (THRE-) Alberta Ltd.  
 PI Wenman WM, Kaul R;  
 DR WPI; 88-39641/48.  
 DR P-PSDB; P80763.  
 PT New antigenic polypeptide cpds. from Chlamydia trachomatis -  
 PT useful in vaccines, and new DNA sequences encoding them.  
 PS Claim 2; Fig 3b; 19pp; English.  
 CC The 18 kD product of the gene is an antigen involved in attachment  
 CC of chlamydial elementary body to the host cell. Antibodies raised  
 CC against the protein can be used for passive immunisation against  
 CC Chlamydial infections.  
 SQ Sequence 1657 BP; 477 A; 319 C; 389 G; 472 T;

Query Match  
 Best Local Similarity 61.5%; Score 16; DB 1; Length 1657;  
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TGCCCTCGAGTGTGAAAGCAAGAT 25  
 DB 901 TTCTCGAGGCGCAACAGAAATAT 878

RESULT 15  
 ID N00005/c  
 AC N00005; standard; cDNA; 99 BP.  
 DT 14-OCT-1992 (first entry)  
 DE Sequence encoding N-terminal of influenza haemagglutinin gene.  
 KW Vaccine; influenza gene; haemagglutinin; antigen; ss.  
 OS Fowl pest virus (Rostock strain).  
 FH Key Location/Qualifiers  
 FT cds 22..99  
 FT /\*tag= a  
 PN BE-882545-A.  
 PD 30-SEP-1980.  
 PF 31-MAR-1980; 882545.  
 PR 02-APR-1979; GB-011487.  
 PA (SEAR) SEARLE G D & CO.  
 DR WPI; 80-73458C/42.  
 DR P-PSDB; P00009.  
 PT Synthetic influenza gene prodn. - from viral RNA by inverse  
 PT transcription then converting DNA to double helix coding for  
 PT antigenic proteins when incorporated in plasmid(s)  
 PS Disclosure; Fig 5; 23pp; French.

Thu Sep 23 10:37:32 1999

us-08-671-757-1.rng

Page 6

CC Viral RNA (VRNA) was isolated from fowl pest virus (Rostock Strain),  
CC subjected to polyadenylation with ATP/poly(A) polymerase, then the  
CC corresp. DNA was synthesised using inverse transcriptase (IT).  
CC N00004 shows the structure of VRNA for influenza haemagglutinin,  
CC indicating the posns. of the start and stop codons. N00005 and  
CC N00006 give the sequences of cDNA corresp. to bps 1-98 and 1028-1129  
CC of N00004 respectively.  
SQ Sequence 99 BP; 32 A; 24 C; 19 G; 24 T;

Query Match 61.5%; Score 16; DB 1; Length 99;  
Best Local Similarity 79.2%; Pred. No. 90;  
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 TGCCTCGAGGTGGAAGCAGAT 25  
DB 57 TGCACCAAGGCGGAAACCAAGAT 34

Search completed: September 22, 1999, 17:00:01  
Job time: 380 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 1999, 16:53:40 ; Search time 307.21 seconds

(without alignments)  
166.941 Million cell updates/sec

Title: US-08-671-757-1

Sequence: 1 ATGCTCGAGGTGCAAAAACAAGATG 26

Scoring table: IDENTITY\_NUC

Searched: 2546578 seqs, 986266752 residues

Database :

EST:  
1: em\_est1:  
2: em\_est2:  
3: em\_est3:  
4: em\_est4:  
5: em\_est5:  
6: em\_est6:  
7: em\_est7:  
8: em\_est8:  
9: em\_est9:  
10: em\_est10:  
11: em\_est11:  
12: em\_est12:  
13: em\_est13:  
14: em\_est14:  
15: em\_est15:  
16: em\_est16:  
17: em\_est17:  
18: em\_est18:  
19: em\_est19:  
20: em\_est20:  
21: em\_est21:  
22: em\_est22:  
23: em\_est23:  
24: em\_est24:  
25: em\_est25:  
26: em\_est26:  
27: em\_est27:  
28: em\_est28:  
29: em\_est29:  
30: em\_est30:  
31: em\_est31:  
32: em\_est32:  
33: em\_est33:  
34: em\_est34:  
35: em\_est35:  
36: em\_est36:  
37: em\_est37:  
38: em\_est38:  
39: em\_est39:  
40: em\_est40:  
41: em\_est41:  
42: em\_est42:  
43: em\_est43:  
44: em\_est44:  
45: em\_est45:  
46: em\_est46:  
47: em\_est47:  
48: em\_est48:  
49: em\_est49:  
50: em\_est50:  
51: em\_est51:  
52: em\_est52:  
53: em\_est53:

54: em\_est22:  
55: em\_est23:  
56: em\_est24:  
57: em\_est25:  
58: em\_est26:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	18.6	71.5	435	22	R23925	R23925 yh28g10.r1
C 2	18	69.2	203	39	AA881975	AA881975 vx31g10.r
C 3	17.6	67.7	524	28	AA116559	AA116559 mp99c11.r
C 4	17.6	67.7	432	30	AA257275	AA257275 MBAFCZ1A7
C 5	17.6	67.7	459	30	AA268482	AA268482 va05d11.r
C 6	17.6	67.7	310	30	AA271856	AA271856 vb73h12.r
C 7	17.6	67.7	432	40	AA981298	AA981298 vx60h11.r
C 8	17.6	67.7	409	47	AT510391	AT510391 mp99c11.y
C 9	17.6	67.7	493	48	AT608046	AT608046 va05d11.y
C 10	17.6	67.7	347	49	AT661202	AT661202 va05d11.x
C 11	17.2	66.2	361	25	N43298	N43298 SW31CA768SK
C 12	17.2	66.2	590	34	AA503372	AA503372 ne44e12.s
C 13	17.2	66.2	612	41	AT047861	AT047861 ud66f06.y
C 14	17.2	66.2	423	23	R85938	R85938 y037906.s1
C 15	17.2	66.2	443	24	H80272	H80272 yu59f01.s1
C 16	17.2	66.2	445	24	H80286	H80286 yu59f01.s1
C 17	17.2	66.2	398	26	W94316	W94316 zh51b12.s1
C 18	17.2	66.2	521	30	AA238826	AA238826 mx92g03.r
C 19	17.2	66.2	567	46	AT425375	AT425375 mx29g03.y
C 20	17.2	66.2	692	49	AT640964	AT640964 AEMTAP28
C 21	16.8	64.6	504	27	AA023634	AA023634 mh80h06.r
C 22	16.8	64.6	521	38	AA777718	AA777718 z304d10.s
C 23	16.8	64.6	406	39	AA871822	AA871822 vq41d09.r
C 24	16.8	64.6	267	50	AV033157	AV033157 AV033157
C 25	16.8	64.6	669	51	AT729324	AT729324 BMLGH131
C 26	16.6	63.8	399	20	Z34559	Z34559 AFTS319.ve
C 27	16.6	63.8	387	26	W77727	W77727 JAO0A005.OA
C 28	16.6	63.8	485	29	AA159799	AA159799 z083a06.s
C 29	16.6	63.8	335	30	AA237987	AA237987 mx29e11.r
C 30	16.6	63.8	485	34	AA533239	AA533239 d153d08.s
C 31	16.6	63.8	170	37	AA711387	AA711387 vt76b11.r
C 32	16.6	63.8	661	38	AA440991	AA440991 LD15709.5
C 33	16.6	63.8	317	38	AA784458	AA784458 e4b02a1.r
C 34	16.6	63.8	279	43	AI210614	AI210614 j9g01a1.r
C 35	16.6	63.8	396	43	AI212866	AI212866 x1f08a1.r
C 36	16.6	63.8	435	43	AI216927	AI216927 M-EST249
C 37	16.6	63.8	574	50	AI655373	AI655373 605010D10
C 38	16.6	63.8	546	50	AI665598	AI665598 605013D12
C 39	16.6	63.8	567	50	AI666167	AI666167 606005G06
C 40	16.6	63.8	588	50	AI667742	AI667742 605026G06
C 41	16.6	63.8	620	50	AI677630	AI677630 605057H12
C 42	16.6	63.8	185	50	AV041189	AV041189 AV041189
C 43	16.6	63.8	582	51	AT727140	AT727140 BMLGH1740

## ALIGNMENTS

RESULT 1  
R23925/c R23925 435 bp mRNA  
LOCUS yh28g10.r1 Soares Placenta Nb2HP Homo sapiens CDNA clone  
DEFINITION IMAGE:131106 5', mRNA sequence.  
ACCESSION R23925  
NID 9778813  
VERSION R23925.1 GI:778813

REFERENCE AUTHORS	TITLE
1 (bases 1 to 544) Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisiel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Thelsting,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and Waterston	The WashU-HHMI Mouse EST Project

Contact: Marra M/Mouse Est Project  
WashU-HHMI Mouse Est Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 28611800

Email: [mouseeest@watson.wustl.edu](mailto:mouseeest@watson.wustl.edu)  
This clone is available royalty-free through LLNL ; contact the  
IMaGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
MGI:352012

Possible reversed clone: similarity on wrong strand  
Seq primer: -28M13 rev2 from Amersham  
High quality sequence stop: 411.

## FEATURES

Location/Qualifiers  
1. .524

```

1      /organism="Mus musculus"
2      /strain="C57BL/6J"
3      /db_xref="taxon:10090"
4      /map="7 q35-q36: 4p16.3; 3; X"
5      /clone="IMAGE:577364"
6      /clone_lib="Soares 2NDMT"
7      /sex="male"
8      /tissue_type="Thymus"
9      /dev_stage="4 weeks"
10     /lab_host="DH10B"
11     /note="Vector: pUT73D-Pac (Pharmacia) with a modified
12     polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
13     TGTACCAACTGTAAGTGGAGCGCCGCGTATTTTATTTTATTTTATTTT
14     3'; double-stranded cDNA was ligated to Eco RI adaptors
15     (Pharmacia), digested with Not I and cloned into the Not I
16     and Eco RI sites of the modified pUT73 vector. RNA
17     extracted by Dr. Bertrand Jordan. Library went through two
18     rounds of normalization, and was constructed by Bento
19     Soares and M.Fatima Bonaldo."
20
21 BASE COUNT      186 a      93 c      108 g      136 t      1 others
22 ORIGIN

```

Query March	67.7%	Score 17.6:	DB 28;	Length 524;
Best Local Similarity	83.3%	Pred No.1.7e+02;		
Matches 20; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;
3 GCCCTGGAGATCGAAAAGCAAGATG	26			
Db 325 GCCCTGGAGATCGAAAAGCAAGATG	348			

RESULT	4
AA257275/c	
LOCUS	
DEFINITION	AA257275            432 bp       mRNA
ACCESSION	M96FCG1ATY3 Brugia malayi adult female cDNA (SAM96BLW-BNAF) Bru
NID	AA257275
VERSION	G1893005
KEYWORDS	AA257275.1     GI:1893005
SOURCE	EST.
ORGANISM	Brugia malayi.
	Brugia malayi.
REFERENCE	Eukaryota; Metazoa; Nematoda; Secernentea; Spirurina; Spiruriida;
AUTHORS	Filarioidea; Onchocercidae; Brugia. Blaxter,M.L., Waterfall,M., Daub,J., Lizotte,M., Baron,L. and Jones,S.J. Genes expressed in adult female Brugia malayi Unpublished (1996) On Sep 12, 1996 this sequence version replaced nt1185443
TITLE	
JOURNAL	
COMMENT	

Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology

University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaetereid.ac.uk  
The ABI trace of this sequence can be viewed at  
<http://www.sanger.ac.uk/brugia/AFC/MBAPC2A7n3.html>  
Seq primer: T3.  
Location/Qualifiers  
1..432

## FEATURES

```

1..432
/organism="Brugia_malay_i"
/db_xref="taxon:6279"
/map="963H04; 6; 6p24"
/clone="AFC21A7"
/clone_lib="Brugia_malay_i_adult_female_CDNA"
(SAW96MIM-fmaf)

```

phoe-vec001; Lambda Uni-zap XR; Site\_1: Eco RI; Site\_2  
Xho I; Lymphatic filarial nematode parasite of humans.  
mRNA was prepared from approximately 50 adult females  
isolated from the peritoneal cavity of jirds and  
converted to double-stranded cDNA using reverse  
transcriptase and oligo(dT). Followed by RNase H and DNA  
pol I. The library has 5 x 10<sup>6</sup> independent recombinants  
and the average insert size is ~900bp. The library was  
constructed by Michelle Lizotte-Waniewski. The  
library is available from Dr.S.A.Williams, email:  
genome@smith.edu "

BASE COUNT	129 a	75 c	101 g	127 t
ORIGIN				

Query March	67.7%	Score 17.6	DB 30	Length 432
Best Local Similarity	83.3%	Pred. No. 1.6e+02		
Matches 20	Conservative 0	Mismatches 4	Indels 0	Gaps 0
QY	2	TGCGTCGAGTGGAAAGACAGAT	25	
Db	323	TGCGTCGAGTGGAAAGACAGAT	300	

LOCUS	AA268482	RESULT 5
DEFINITION	AA268482 459 bp mRNA	EST
ACCESSION	AA268482	21-MAR-1997
NID	IMAGE:1722037 5', mRNA sequence.	
KEYWORDS	AA268482.1 GI:1905226	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
REFERENCE	1 (bases 1 to 459)	
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque, Geisels,S., Kucab,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.	
TITLE	The WashU-HHMI Mouse EST Project	
JOURNAL	Unpublished (1996)	
COMMENT	On Sep 12, 1996 this sequence version replaced gi:1404853	

Contact: Marra M/Mouse EST Project  
WashU-HHMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: mousestewatson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:447533

Seq primer: -28m13 rev2 ET from Amersham  
High quality sequence stop: 453.

## FEATURES

## source

Location/Qualifiers  
1. .459  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:722037"  
/clone\_lib="Soares mouse lymph node NbMLN"  
/sex="male"  
/tissue\_type="lymph node"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia)  
with a modified polylinker; Site.1: Not I; Site.2: Eco RI;  
1st strand cDNA was primed with a Not I - oligo(dT) primer  
[5',  
TGTACCAATCTGAAGTGGAGCGCGCCGACTTTTCTTTTCTTTTCTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. RNA  
provided by Dr. Bertrand Jordan. Library constructed and  
normalized by Bento Soares and M. Fatima Bonaldo."

## BASE COUNT

107 a 126 c 110 g 116 t

## ORIGIN

Query Match 67.7%; Score 17.6; DB 30; Length 459;  
Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GCCTCGAGTCGAAAGCAAGATG 26

Db 197 GCCTCGAGTCGAAATCAAGATG 220

RESULT 6  
AA271856 310 bp mRNA EST 26-MAR-1997  
LOCUS v073h12.r1 Soares mouse 3NM12.5 Mus musculus cDNA clone  
DEFINITION IMAGE:762695.5 similar to SW:RS27\_HUMAN P42677.405 RIBOSOMAL  
PROTEIN S27; mRNA sequence.  
ACCESSION AA271856  
NID 91910185  
VERSION AA271856.1 GI:1910185  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 310)  
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.  
TITLE The WashU-HMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT On Sep 12, 1996 this sequence version replaced gi:1402102.

Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mousestewatson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:463615  
putative full length read  
vector to vector length is 320  
Seq primer: -28m13 rev2 ET from Amersham.

## FEATURES

## source

Location/Qualifiers  
1. .310  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:762695"  
/clone\_lib="Soares mouse 3NM12.5"  
/sex="unknown"  
/tissue\_type="fetus"  
/dev\_stage="12.5dpc total fetus"  
/lab\_host="DH10B"  
/note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia)  
with a modified polylinker; Site.1: Not I; Site.2: Eco RI;  
1st strand cDNA was primed with a Not I - oligo(dT) primer  
[5',  
TGTACCAATCTGAAGTGGAGCGCGCCGCTTTTCTTTTCTTTTCTTTT  
3']; on total mouse RNA (provided by Minoru Ko, Wayne  
State Univ.); double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT7T3 vector.  
Library went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

## BASE COUNT

89 a 80 c 73 g 68 t

## ORIGIN

Query Match 67.7%; Score 17.6; DB 30; Length 310;  
Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GCCTCGAGTCGAAAGCAAGATG 26

Db 216 GCCTACAGTCGAAACCAAGTCG 239

RESULT 7  
AA981298 432 bp mRNA EST 27-MAY-1998  
LOCUS vx60h11.r1 Stratiogene mouse macrophage (#937306) Mus musculus cDNA  
DEFINITION IMAGE:1279653.5 similar to TR:014730.014730 HOMOLOG OF THE  
ASPERKILIN NIDULANS SUDG GENE PRODUCT.; mRNA sequence.  
ACCESSION AA981298  
NID 93159834  
VERSION AA981298.1 GI:3159834  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 432)  
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.  
TITLE The WashU-HMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT On Jan 19, 1998 this sequence version replaced gi:2285238.

Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mousestewatson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:671453  
Possible reversed clone: similarity on wrong strand



Seq primer: -28ml3 rev1 ET from Amersham  
High quality sequence stop: 402.

## FEATURES

Location/Qualifiers  
1..432

ORIGIN  
148 a 76 c 100 g 108 t  
BASE COUNT  
148 a 76 c 100 g 108 t  
ORIGIN  
148 a 76 c 100 g 108 t

Query Match  
Best Local Similarity 83.3%; Score 17.6; DB 40; Length 432;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
DB 296 GCCTGAGAGTCGAAAGCAAGATG 319

## RESULT 8

LOCUS AI510391 409 bp mRNA EST 12-MAR-1999  
DEFINITION mp96cl1.y1 Soares 2MBMT Mus musculus CDNA IMAGE:577364 5' similar to TR:014730 014730 HOMOLOG OF THE ASPERGILLUS NIDULANS SUDG GENE PRODUCT. ; mRNA sequence.  
ACCESSION AI510391  
VERSION 94409296  
KEYWORDS AI510391.1 GI:4409296  
SOURCE EST.  
ORGANISM house mouse.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.

## REFERENCE

AUTHORS Maria M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.  
The WashU-NCI Mouse EST Project 1999  
Unpublished (1999)

TITLE JOURNAL  
COMMENT On Jun 5, 1998 this sequence version replaced g1:3189086.

## FEATURES

source  
1..409  
Location/Qualifiers

ORIGIN  
148 a 76 c 100 g 108 t  
BASE COUNT  
148 a 76 c 100 g 108 t  
ORIGIN  
148 a 76 c 100 g 108 t

Query Match  
Best Local Similarity 83.3%; Score 17.6; DB 47; Length 409;  
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
DB 325 GCCTGAGAGTCGAAAGCAAGATG 348

## RESULT 9

LOCUS AI608046 493 bp mRNA EST 21-APR-1999  
DEFINITION va05d11.y1 Soares mouse lymph node NbMLN Mus musculus CDNA clone IMAGE:722037 5', mRNA sequence.  
ACCESSION AI608046  
VERSION 94617213  
KEYWORDS AI608046.1 GI:4617213  
SOURCE EST.  
ORGANISM house mouse.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.

## REFERENCE

AUTHORS Maria M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.  
The WashU-NCI Mouse EST Project 1999  
Unpublished (1999)

TITLE JOURNAL  
COMMENT On Jun 22, 1998 this sequence version replaced g1:3246994.

## FEATURES

source  
1..493  
Location/Qualifiers

ORIGIN  
145 a 73 c 82 g 109 t  
BASE COUNT  
145 a 73 c 82 g 109 t  
ORIGIN  
145 a 73 c 82 g 109 t

```

/clone_lib="Soares mouse lymph node NbXLN"
/sex="male"
/tissue_type="lymph node"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: pT713D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5',
TGTTCACATCTGAGATGGAGCGCGCCGCATCTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT713 vector. RNA
provided by Dr. Bertrand Jordan. Library constructed and
normalized by Bento Soares and M. Fatima Ronaldo."

```

Query Match	67.7%	Score 17.6;	DB 48;	Length 493;
Best Local Similarity	83.3%	Pred. No. 1.7e+02;		
Matches 20;	Conservative	0;	Mismatches 4;	Indels 0;
				Gaps 0;

Qy            3 GCCTGAGGTGCAAAAGCAAGATG    26  
               | | | | | | | | | |  
Db          198 GGCTGGAGGTCCAAATCAAGATG    221

RESULT	10
AI661202/c	
LOCUS	AI661202 347 bp mRNA EST 10-MAY-1999
DEFINITION	W050511.x1 Soares mouse lymph node NbM1.N Mus musculus cDNA clone IMAGE:722037 3, mRNA sequence.

ACCESSION	A1661202
NUM	947664785
VERSION	A1661202.1
KEYWORDS	GT.4764785
SOURCE	EST.
ORGANISM	house mouse. Mus musculus

TITLE The WashU-NCI Mouse EST Project 1999  
 JOURNAL Unpublished (1999)  
 COMMENT On Mar 20, 1998 this sequence version replaced gi:2960279.

Other\_ESTs: va0a011.yi  
Contact: Maria M/MashU-MCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LINTL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MG1.447533  
This clone was previously sequenced on the 5' end only, this new  
data is from the 3' end  
High quality sequence stop: 328.  
Location/Qualifiers  
FEATURES  
source  
1..347

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/map="19p12.00:19q13.2"
/clone_image="722037"
/clone_lib="Scotars mouse lymph node NBLMLN"
/sex="male"
/tissue_type="lymph node"

```

```

/dev stage="4 weeks"
/lab host="DH10B"
/notes="Organ: lymph node; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker. Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5',
TGTTCACATCTGAAGTGGAGGCGGCGGACATCTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library constructed and
normalized by Bento Soares and M.Fátima Bonaldo."
BASE COUNT      104 a      75 c      88 g      80 t
ORIGIN
Query Match      67.7%; Score 17.6; DB 49; Length 347;
      Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

Query Match	67.7%	Score 17.6	DB 49	Length 347
Best Local Similarity	83.3%	Pred. No. 1.6e+02		
Matches 20	Conservative 0	Mismatches 4	Indels 0	Gaps 0
QY	3	GCTTCGAGGTCGAAAAGCAAGATG	26	
		11111111111111111111111111		
Db	329	GCTGCGAGGTCGAAATCAAGATG	306	

RESULT	11
MA3298	
LOCUS	MA3298 361 bp mRNA
DEFINITION	Brugia malayi infective larva cDNA (SAB447L-Bml3)
ACCESSION	Brugia malayi cDNA clone SM3ICA768 5' mRNA sequence.
MA3298	
NID	g1167680

VERSION	M43298.1	GI:1167680
KEYWORDS	EST.	
SOURCE	Brugia malayi.	
ORGANISM	Brugia malayi	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Nematoda; Secernentea; Spirurina; Spiruridae;  
Filarioidea; Onchocercidae; Brugia  
1 (bases 1 to 361)  
Blaxter, M.L., Raghavan, N., Ghosh, I., Guiliano, D., Lu, W.,  
Williams, S.S., Slatko, B. and Scott, A.L.  
Genes expressed in Brugia malayi infective third stage larvae  
Mot. Biochem. Parasitol. 77 (1), 77-94 (1996)  
On Apr 11, 1993 this sequence version replaced gi:692807.

Contact: Steven A. Williams  
 Molecular Parasitology  
 Smith College Department of Biological Sciences  
 Department of Biological Sciences, Clark Science Center, Smith  
 College, Northampton, MA, 01063, USA  
 Tel: 4135853826  
 Fax: 4135853786  
 Email: [genome@smith.edu](mailto:genome@smith.edu)  
 Seq primer: pBluescript SK  
 High quality sequence stop: 433.  
 Location/Qualifiers  
     I..361

```

/organism="Brugia malayi"
/strain="JRS Labs"
/db_xref="taxon:6279"
/clone="SM3ICAN768"
/clone_lib="Brugia malayi infective larva cDNA
(SAM94WL-Bml3)"
/lab_host="X11-Blue MRF"
/note="vector: lambda unizap XR; Site_1: EcoR I; Site_2:
Xho I; Lymphatic filarial nematode parasite of humans.
mRNA was prepared from third stage infective larvae of
Brugia malayi isolated from mosquitoes 10 days after
infection and converted to double stranded cDNA using
reverse transcriptase and oligo(dT) followed by RNase H
and DNAPol I. The library had 1.6 x 10E6 independent
recombinants and average insert size was 900 base pairs.
The library was constructed by Wenhong Lu. The library is

```



Matches 19: Conservative 0: Mismatches 3: Indels 0: Gaps 0:

QY 5 CTCGAGTCGAAAGCAGATG 26

Db 574 CCCGAGGTGGAAGCAAGAG 595

#### RESULT 14

R85938/c

LOCUS R85938 423 bp mRNA EST 14-AUG-1995  
DEFINITION y037906.s1 Soares adult brain N2b4HB55Y Homo sapiens cDNA clone  
IMAGE:180154 3', mRNA sequence.

#### ACCESSION

R85938

NID 9944344

VERSION R85938.1

KEYWORDS GI:944344

SOURCE EST.

ORGANISM human.

REFERENCE Homo sapiens

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 423)  
Hillier, L., Clark, N., Dubuque, T., Eliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.  
The WashU-Merck EST Project  
Unpublished (1995)  
On Nov 4, 1993 this sequence version replaced.

TITLE The WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT On Nov 4, 1993 this sequence version replaced.

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 1422  
High quality sequence stops: 320  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert Length: 1422 Std Error: 0.00  
Seq primer: Promega -21m13  
High quality sequence stop: 320.

#### FEATURES

source

1. .423

/organism="Homo sapiens"

/db\_xref="GDB:382679"

/db\_xref="taxon:9606"

/clone="IMAGE:180154"

/clone\_lib="Soares adult brain N2b4HB55Y"

/sex="Male"

/dev\_stage="55-year old"

/lab\_host="DH10B (ampicillin resistant)"

/note="Organ: brain; Vector: p773D (Pharmacia) with a  
modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TGTTCACATCTGAACTGGAGCGCCGCTTTTCTTTTCTTTT 3'],  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified p773 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 53. Library constructed by Bento  
Soares and M.Fatima Bonaldo. The adult brain RNA was  
provided by Dr. Donald H. Gilden. Tissue was acquired  
17-18 hours after death which occurred in consequence of a  
ruptured aortic aneurysm. RNA was prepared from a pool of  
tissues representing the following areas of the brain:  
frontal, parietal, temporal and occipital cortex from the  
left and right hemispheres, subcortical white matter,  
basal ganglia, thalamus, cerebellum, midbrain, pons and  
medulla."

BASE COUNT

93 a 118 c 121 g 88 t 3 others

#### ORIGIN

Query Match 65.4%; Score 17; DB 23; Length 423;  
Best Local Similarity 80.0%; Pred. No. 3.1e+02;  
Matches 20: Conservative 0: Mismatches 5: Indels 0: Gaps 0:

QY 2 TCCCTGAGTCGAAAGCAGATG 26

Db 419 TTCACCGAGCGCCGAAAGCAGATG 395

#### RESULT 15

H80272/c

LOCUS H80272 443 bp mRNA EST 09-NOV-1995  
DEFINITION y059601.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
IMAGE:230401 3', mRNA sequence.

#### ACCESSION

H80272

NID 91058361

VERSION H80272.1

KEYWORDS GI:1058361

SOURCE EST.

ORGANISM human.

REFERENCE Homo sapiens

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 443)  
Hillier, L., Clark, N., Dubuque, T., Eliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.  
The WashU-Merck EST Project  
Unpublished (1995)  
On Apr 14, 1993 this sequence version replaced gi:837779.

TITLE The WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT On Apr 14, 1993 this sequence version replaced gi:837779.

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 747  
High quality sequence stops: 329  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert Length: 747 Std Error: 0.00  
Seq primer: Promega -21m13  
High quality sequence stop: 329.

#### FEATURES

source

1. .443

/organism="Homo sapiens"

/db\_xref="GDB:3781497"

/db\_xref="taxon:9606"

/clone="IMAGE:230401"

/clone\_lib="Soares fetal liver spleen INFLS"

/sex="Male"

/dev\_stage="20 week-post conception fetus"

/lab\_host="DH10B (ampicillin resistant)"

/note="Organ: liver and spleen; Vector: p773D (Pharmacia)  
with a modified polylinker; Site:1: Pac I; Site:2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5' AACGAGAGATTAATTAAGATCTTTTCTTTTCTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adapters  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified p773 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT

122 a 94 c 97 g 128 t 2 others

Query Match

65.4%; Score 17; DB 24; Length 443;

Thu Sep 23 10:37:34 1999

us-08-671-757-1.rst

Page 9

Best Local Similarity 80.0%; Pred No. 3.1e+02;  
Matches 20; Conservative 0; Mismatches 5; Indels 0;

QY 2 TGGCTGAGGTGGAAGCAGATG 26  
||| | ||||| ||||| ||| |  
Db 206 TGGGTGAGGTCCAAAGGAGGTTG 182

Search completed: September 22, 1999, 16:59:01  
Job time: 321 sec

